

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

**Defective images within this document are accurate representation of
The original documents submitted by the applicant.**

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

(19)日本国特許庁(JP)

(12)特 許 公 報 (B2)

(11)特許出願公告番号

特公平6-54440

(24)(44)公告日 平成6年(1994)7月20日

(51)Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 3 1 Z	9379-5H		
	5 2 1 C	9379-5H		
5/06	D	9379-5H		

発明の数1(全 3 頁)

(21)出願番号	特願昭61-156409	(71)出願人	999999999 日本電気株式会社 東京都港区芝5丁目7番1号
(22)出願日	昭和61年(1986)7月2日	(72)発明者	渡辺 隆夫 東京都港区芝5丁目33番1号 日本電気株 式会社内
(65)公開番号	特開昭63-11998	(74)代理人	弁理士 京本 直樹 (外2名)
(43)公開日	昭和63年(1988)1月19日		審査官 橋本 武

(54)【発明の名称】 音声認識におけるマッチング方法

1

【特許請求の範囲】

【請求項1】 標識パタンとのパタンマッチングにより入力された音声認識する音声認識におけるマッチング方法において、音声を構成するユニットごとに複数個のユニット標準パタンを有し前記複数個のユニット標準パタンから選択されたものを連結することにより得られたパタンと入力されたパタンとのマッチングを行い、かつ前記ユニット標準パタンの選択は入力パタンとユニット標準パタンとを予備的に比較することによって行なうものであることを特徴とする音声認識におけるマッチング方法。

【発明の詳細な説明】

(産業上の利用分野)

本発明は音声認識におけるマッチング方法に関し、特に、パタンマッチングにより認識を行う技術の改良を図

2

った音声認識におけるマッチング方法に関する。

(従来の技術)

音声認識の方法として、認識すべきパタンをあらかじめ標準パタンとして登録しておき、認識時に入力される未知パタンを標準パタンと比較し最も類似性の高いものを認識結果として決定するパタンマッチング法は従来から広く用いられている。

このような標準パタンによる方法は、認識対象のカテゴリ数が増加するに従って標準パタンの登録の負担が増加し、同時に標準パタンを保持するための記憶量も増加する。このような問題を解決するため、音声をより細かなユニットの結合とみなし、標準パタンをユニット毎に用意しこれらユニット標準パタンを連結することにより標準パタンを生成する方法が試みられている。

音節はこのようなユニットとして典型的なものであり、

日本語の場合には約100種類の音節の連結により任意の音声の生成することが可能である。

パターンマッチング法では、認識精度を向上させるため、各カテゴリ1個の標準パターンだけでなく複数の標準パターンを用いることが一般に有効である。しかしながら、上記のようなユニット標準パターンを用いる場合には、ユニット標準パターンの個数を増やすと連結のしかたは著しく増加する。例えば「カワサキ」という単語を考えた場合、各音節ごとに5個の音節標準パターン（ユニット標準パターン）を用意しただけでも連結のしかたは5通りとなりマッチングの処理量は膨大なものとなる。連結単語音声認識で用いられている2段DP Dynamic Programming, 動的計画法）マッチング法と呼ばれる方法を適用すると（この場合各音節が単語とみなされる）、この処理量は大幅に削減されるがそれでも処理量は依然として大きい。なお、2段DPマッチング法の詳細アルゴリズムに関しては“日経エレクトロニクス, 昭58, No. 11, Vol. 4, PP. 171-207”に詳述されている。

（発明が解決しようとする問題点）

上述した如く従来のこの種の音声認識におけるマッチング方法は、ユニット標準パターンの個数に対応して連結数が著しく増加しパターンマッチングにおける処理量が膨大化することが避けられないという欠点がある。

本発明の目的は上述した欠点を除去し、複数のユニット標準パターンを利用しても処理量の増加をもたらすことのない音声認識におけるマッチング方法を提供することにある。

（発明の構成）

本発明の方法は、標準パターンとのパターンマッチングにより入力された音声の認識する音声認識におけるマッチング方法において、音声を構成するユニットごとに複数の標準パターンを有し前記複数のユニット標準パターンから選択されたものを連結することにより得られたパターンと入力されたパターンとのマッチングを行い、かつ前記ユニット標準パターンの選択は入力パターンとユニット標準パターンとを予備的に比較することによって行われる手段を備えて構成される。

（作用）

本発明の作用を、単語音声認識において標準パターンのユニットを音節とした場合を例にとって説明する。

入力された単語音声パターンは、音声分析(FFT分析, LPC分析等)により特徴ベクトルの系列 $A = \{a(j), j = 1, \dots, J\}$ と表わされるものとする。また認識対象となる単語はおの音節シンボルの系列として表わされるものとする。これを $S = \{s(p), p = 1, \dots, P\}$ と表わす。

音節標準パターンは認識に先立って登録されており、入力パターンと同様に特徴ベクトルの系列として表わされるものとする。これを $B^{(k)} = \{b^{(k)}(i), i = 1, \dots,$

$1\}$ と表わす。 $B^{(k)}$ は音節 m の第 k 番目の標準パターン

を表わす。

問題は入力パターン A と音節シンボル系列 S との類似度を求めることである。各音節の標準パターン数 $K(m)$ が1であるときは音節の標準パターンを音節シンボル系列の順に連結することによって単語標準パターンを生成することが容易にできる。

この場合には、代表的なパターンマッチング法として広く実用に供されている動的計画法を利用した時間軸正規化マッチング法、いわゆるDPマッチング法（“日本音響学会誌, Vol. 27, No. 9, P. 483”）を適用することができる。

$K(m)$ が2以上の場合には、上記のようにして生成する単語標準パターンは複数個となり単語マッチングの処理量が増加する。これに対し各音節毎に $K(m)$ 個の音節標準パターンのうちから最も入力パターンに類似していると思われるものをとり出してこれらを連結することにすれば、生成される単語標準パターンは1つであり、処理量は大幅に減少できる。

このような方法としては、入力音声のうち、ある音節の存在する可能性のある区間を適当に切り出し、この区間で各音節標準パターンとの類似度を求める方法いわゆる音節セグメンテーション(Segmentation)がある。

第2図は本発明における音節セグメンテーションの原理図である。第2図は、例として2音節長の単語「りく」の単語標準パターンを生成する場合をとり、入力パターンの区間 A と「り」の標準パターンとの類似度および区間 B と「く」の標準パターンとの類似度をそれぞれ求めればよいことを示している。このとき類似度として2つのパターンの包含関係を表わす値、例えば、特願昭60-153217に示される方法において用いられている値 Δ を用いることができる。

区間 A, B を求めるには、振巾パターンの特徴を利用する方法や単純に入力区間を等分する方法など任意のものが可能である。第1図では、入力区間 $[S, E]$ を2つに等分する点 X を求め、適当な許容限度 ϵ を用いて、 $A = [S, X + \epsilon]$ $B = [X - \epsilon, E]$ として与えている。

〔実施例〕

次に図面を参照して本発明を詳細に説明する。

第1図は本発明の一実施例を示すブロック図である。第1図において、入力パターンバッファ1の内容は単語辞書5の各単語とのパターンマッチングが行われる。音節セグメンテーション処理部2は、バッファ1の入力パターンを単語辞書5に記憶されている単語を仮説して各音節に対応する区間を求める。音節標準パターン選択部3では、処理部2で求められた各音節区間のパターンについて、標準パターンバッファ6中の各音節標準パターン（但し音節名が、仮説された単語内の対応する音節と一致するもの）との比較を行い最も類似している音節標準パターンを連結標準パターンバッファ7へ出力する。連結標準パターンバッファ7には、選択部3で決定された音節標準パターンを連

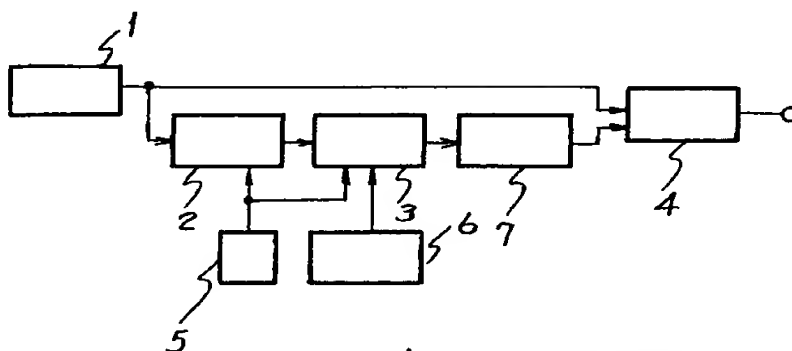
結したパタンが単語標準パタンとして格納される。

単語マッチング部4は、パタンバッファ1上の入力パタンとパタンバッファ7上の単語標準パタンとのマッチングを行い類似度を出力する。

以上の例において、セグメンテーション処理部2は、単純に音節数で等分したもの、単語に固有の情報（無声部を含むエネルギーの山がある等）を利用したもの等任意のものが使用可能である。標準パタン選択部3も、入力パタンの部分パタンと標準パタンの類似性を判定できるものであれば任意のものが可能であるが、効率の改善が本発明の利点であるので計算の容易なものである程よい。特願昭60-153217で述べられている方法はその一例である。単語マッチング部4も任意のものが可能であるが、DPマッチング法が認識精度の点で優れていることはよく知られており、この場合には最も適切である。

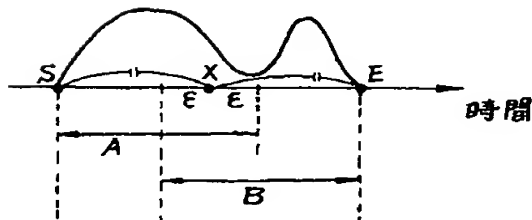
*

【第1図】



- 1 ……入力パタンバッファ
- 2 ……セグメンテーション処理部
- 3 ……標準パタン選択部
- 4 ……単語マッチング部
- 5 ……単語辞書
- 6 ……標準パタンバッファ
- 7 ……連結標準パタンバッファ

【第2図】



Translation of JP, 6-54440, B2

[Title of the Invention] Matching Method in Speech Recognition

5 [Scope of Claim for Patent]

[Claim 1] A matching method in speech recognition for recognizing an input speech through pattern matching with an indicator pattern by providing a plurality of unit standard patterns every unit forming said speech for matching a pattern
10 obtained by connecting those selected from said plurality of unit standard patterns with an input pattern while selecting said unit standard patterns by preliminarily comparing said input pattern with said unit standard patterns.

[Detailed Description of the Invention]

15 (Industrial Field of Application)

The present invention relates to a matching method in speech recognition, and more particularly, it relates to a matching method in speech recognition aiming at improving a technique for making recognition by pattern matching.

20 (Prior Art)

As a method of speech recognition, a pattern matching method of previously registering a pattern to be recognized as a standard pattern for comparing unknown patterns input in recognition with the standard pattern and deciding the most
25 similar one as the result of recognition is widely employed

in general.

In such a method employing the standard pattern, the burden of registration of standard patterns is increased as the number of categories of objects to be recognized is increased, while the storage capacity for holding the standard patterns is also increased at the same time. In order to solve this problem, a method of recognizing a speech as connection of finer units, preparing standard patterns for the respective units and connecting these unit standard patterns with each other thereby forming a standard pattern is attempted.

Syllables are typical as such units, and arbitrary speeches can be formed by connecting about 100 types of syllables in the case of Japanese.

In the pattern matching method, it is generally effective to employ not only a single standard pattern but also a plurality of standard patterns for each category, in order to improve the precision of recognition. In the case of employing the aforementioned unit standard patterns, however, the number of manners of connecting the patterns is remarkably increased if the number of the unit standard patterns is increased. Considering a word "Kawasaki", for example, the number of manners of connecting patterns reaches 5^4 to result in an enormous quantity of matching processing when five syllable standard patterns (unit standard patterns) are simply prepared for each syllable. While this quantity of processing is

remarkably reduced when a method called a two-stage DP (dynamic programming) matching method employed in connective speech recognition is applied, the quantity of processing is still large. The algorithm of the two-stage DP matching method is described in detail in "Nikkei Electronics", No. 1, 1983, Vol. 4, pp. 171 to 207.

(Problem to be Solved by the Invention)

In the conventional matching method for this type of speech recognition, as hereinabove described, the number of connection is remarkably increased in correspondence to the number of unit standard patterns to unavoidably result in an enormous quantity of processing in pattern matching.

In order to solve the aforementioned problem, an object of the present invention is to provide a matching method in speech recognition not increasing the quantity of processing also when utilizing a plurality of unit standard patterns.

(Structure of the Invention)

According to the present invention, a matching method in speech recognition for recognizing an input speech through pattern matching with an indicator pattern comprises means providing a plurality of unit standard patterns every unit forming the speech for matching a pattern obtained by connecting those selected from the said plurality of unit standard patterns with an input pattern while selecting the said unit standard patterns by preliminarily comparing the

input pattern with the unit standard patterns.

(Function)

The function of the present invention is described with reference to a case of forming units of standard patterns by syllables in word recognition.

It is assumed that an input word pattern is expressed as a series $A = \{a(j), j = 1, \dots, J\}$ of feature vectors by speech analysis (FFT analysis, LPC analysis or the like). It is also assumed that the word to be recognized is expressed as a series of syllable symbols. This is expressed as $S = \{s(p), p = 1, \dots, P\}$.

It is assumed that syllable standard patterns are registered in advance of recognition and expressed as a series of feature vectors similarly to the input pattern. This is expressed as $B^{m,k} \{b^{m,k}(i), i = 1, \dots, I\}$. $B^{m,k}$ represents a k-th standard pattern of a syllable m.

The point is to obtain the degree of similarity between the input pattern A and the syllable symbol series S. When the number $K(m)$ of standard patterns of each syllable is 1, a word standard pattern can be readily formed by connecting the standard patterns of the syllables in order of the syllable symbol series.

In this case, it is possible to apply a time axis normalization matching method, the so-called DP matching method (Journal of the Acoustical Society of Japan, Vol. 27,

No. 9, P. 483) employing the dynamic programming widely put into practice as a typical pattern matching method.

When the number $K(m)$ is at least 2, a plurality of word standard patterns can be formed in the aforementioned manner, to increase the quantity of processing of word matching. When extracting that conceivably most similar to the input pattern from $K(m)$ syllable standard patterns every syllable and connecting such patterns with each other, a single word standard pattern is formed and the quantity of processing can be remarkably reduced.

Such a method includes the so-called syllable segmentation of properly slicing a section having a possibility of including a certain syllable from an input speech and obtaining the degree of similarity to each syllable standard pattern in this section.

Fig. 2 is a principle diagram of syllable segmentation in the present invention. With reference to a case of forming word standard patterns of a word "riku" having two syllables, for example, Fig. 2 shows that the degree of similarity between a section A of an input pattern and a standard pattern of "ri" and the degree of similarity between a section B and a standard pattern of "ku" may be obtained respectively. In this case, a value expressing the inclusion relation between the two patterns, such as a value Δ employed in a method described in Japanese Patent Application No. 60-153217 (1985), for example,

can be employed.

The sections A and B can be arbitrarily obtained by a method utilizing features of amplitude patterns or a method simply equally dividing the input section. Referring to Fig. 2, a point X equally dividing the input section [S, E] into two sections is obtained for giving the sections as $A = [S, X + \varepsilon]$ and $B = [X - \varepsilon, E]$ with a proper tolerance limit ε .

[Embodiment]

The present invention is now described in detail with reference to Fig. 1.

Fig. 1 is a block diagram showing an embodiment of the present invention. Referring to Fig. 1, the contents of an input pattern buffer 1 are subjected to pattern matching with respective words of a word dictionary 5. A syllable segmentation processing part 2 hypothesizes any word stored in the word dictionary 5 with reference to an input pattern of the buffer 1 and obtains sections corresponding to respective syllables. A syllable standard pattern selection part 3 compares the patterns of the respective syllable sections obtained by the processing part 2 with respective syllable standard patterns (those having syllable names matching with the corresponding syllables in the hypothesized word) in a standard pattern buffer 6 and outputs the most similar syllable standard patterns to a connected standard pattern buffer 7. The connected standard pattern buffer 7 stores a

pattern obtained by connecting the syllable standard patterns decided by the selection part 3 as a word standard pattern. A word matching part 4 matches the input pattern on the pattern buffer 1 with the word standard pattern on the pattern buffer 5 7 and outputs the degree of similarity.

In the aforementioned embodiment, the segmentation processing part 2 can be formed by an arbitrary one such as that simply equally dividing a word in response to the number of syllables, that simply utilizing specific information 10 (having a peak of energy including a silent part etc.) or the like. The standard pattern selection part 3 can also be formed by an arbitrary one so far as the same can determine similarity between a partial pattern of an input pattern and a standard pattern, while the advantage of the present invention resides 15 in improvement of efficiency and hence that implementing simple calculation is preferable. The method described in Japanese Patent Application No. 60-153217 is an example thereof. The word matching part 4 can also be formed by an arbitrary one, while it is well known that the DP matching method is superior 20 in precision of recognition, and this is most appropriate in this case.

(Effect of the Invention)

According to the present invention, as hereinabove described, matching can be performed without increasing the 25 quantity of processing while employing a plurality of unit

standard patterns, whereby a matching method in speech recognition capable of implementing pattern matching remarkably improving precision of recognition and efficiency can be effectively implemented.

5 [Brief Description of the Drawings]

Fig. 1 is a block diagram showing an embodiment of the present invention, and Fig. 2 is a principle diagram of syllable segmentation in the present invention.

10 1 ... input pattern buffer, 2 ... segmentation processing part, 3 ... standard pattern selection part, 4 ... word matching part, 5 ... word dictionary, 6 ... standard pattern buffer, 7 ... connected standard pattern buffer.

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許出願公告番号

特公平6-54440

(24)(44)公告日 平成6年(1994)7月20日

(51)Int.Cl. ⁴	識別記号	庁内整理番号	FI	技術表示箇所
G10L 3/00	531 Z	9379-5H		
	521 C	9379-5H		
5/06	D	9379-5H		

発明の数1(全 3 頁)

(21)出願番号 特願昭61-156409

(22)出願日 昭和61年(1986)7月2日

(65)公開番号 特開昭63-11998

(43)公開日 昭和63年(1988)1月19日

(71)出願人 999999999

日本電気株式会社

東京都港区芝5丁目7番1号

(72)発明者 渡辺 隆夫

東京都港区芝5丁目33番1号 日本電気株式会社内

(74)代理人 弁理士 京本 直樹 (外2名)

審査官 橋本 武

(54)【発明の名称】 音声認識におけるマッチング方法

1

【特許請求の範囲】

【請求項1】 標識ボタンとのボタンマッチングにより入力された音声进行認識する音声認識におけるマッチング方法において、音声を構成するユニットごとに複数個のユニット標準ボタンを有し前記複数個のユニット標準ボタンから選択されたものを連結することにより得られたボタンと入力されたボタンとのマッチングを行い、かつ前記ユニット標準ボタンの選択は入力ボタンとユニット標準ボタンとを予備的に比較することによって行なうものであることを特徴とする音声認識におけるマッチング方法。

【発明の詳細な説明】

(産業上の利用分野)

本発明は音声認識におけるマッチング方法に関し、特に、ボタンマッチングにより認識を行う技術の改良を図

2

った音声認識におけるマッチング方法に関する。

(従来の技術)

音声認識の方法として、認識すべきボタンをあらかじめ標準ボタンとして登録しておき、認識時に入力される未知ボタンを標準ボタンと比較し最も類似性の高いものを認識結果として決定するボタンマッチング法は従来から広く用いられている。

このような標準ボタンによる方法は、認識対象のカテゴリ数が増加するに従って標準ボタンの登録の負担が増加し、同時に標準ボタンを保持するための記憶量も増加する。このような問題を解決するため、音声をより細かなユニットの結合とみなし、標準ボタンをユニット毎に用意しこれらユニット標準ボタンを連結することにより標準ボタンを生成する方法が試みられている。

音節はこのようなユニットとして典型的なものであり、

日本語の場合には約100種類の音節の連結により任意の音声を生成することが可能である。

ボタンマッチング法では、認識精度を向上させるため、各カテゴリ1個の標準ボタンだけでなく複数の標準ボタンを用いることが一般に有効である。しかしながら、上記のようなユニット標準ボタンを用いる場合には、ユニット標準ボタンの個数を増やすと連結のしかたは著しく増加する。例えば「カワサキ」という単語を考えた場合、各音節ごとに5個の音節標準ボタン(ユニット標準ボタン)を用意しただけでも連結のしかたは5⁴通りとなりマッチングの処理量は膨大なものとなる。連結単語音声認識で用いられている2段DP Dynamic Programming, 動的計画法) マッチング法と呼ばれる方法を適用すると(この場合各音節が単語とみなされる)、この処理量は大幅に削減されるがそれでも処理量は依然として大きい。なお、2段DPマッチング法の詳細アルゴリズムに関しては「日経エレクトロニクス、昭58、No. 1, Vol. 4, PP. 171-207」に詳述されている。

(発明が解決しようとする問題点)

上述した如く従来のこの種の音声認識におけるマッチング方法は、ユニット標準ボタンの個数に対応して連結数が著しく増加しボタンマッチングにおける処理量が膨大化することが避けられないという欠点がある。本発明の目的は上述した欠点を除去し、複数のユニット標準ボタンを利用しても処理量の増加をもたらすことのない音声認識におけるマッチング方法を提供することにある。

(発明の構成)

本発明の方法は、標準ボタンとのボタンマッチングにより入力された音声認識する音声認識におけるマッチング方法において、音声を構成するユニットごとに複数の標準ボタンを有し前記複数のユニット標準ボタンから選択されたものを連結することにより得られたボタンと入力されたボタンとのマッチングを行い、かつ前記ユニット標準ボタンの選択は入力ボタンとユニット標準ボタンとを予備的に比較することによって行われる手段を備えて構成される。

(作用)

本発明の作用を、単語音声認識において標準ボタンのユニットを音節とした場合を例にとりて説明する。

入力された単語音声ボタンは、音声分析(FFT分析、LPC分析等)により特徴ベクトルの系列 $A = \{a(j), j = 1, \dots, J\}$ と表わされるものとする。また認識対象となる単語はおの音節シンボルの系列として表わされるものとする。これを $S = \{s(p), p = 1, \dots, P\}$ と表わす。

音節標準ボタンは認識に先立って登録されており、入力ボタンと同様に特徴ベクトルの系列として表わされるものとする。これを $B^{m,k} = \{b^{m,k}(i), i = 1, \dots, I\}$ と表わす。 $B^{m,k}$ は音節 m の第 k 番目の標準ボタン

を表わす。

問題は入力ボタン A と音節シンボル系列 S との類似度を求めることである。各音節の標準ボタン数 $K(m)$ が1であるときは音節の標準ボタンを音節シンボル系列の順に連結することによって単語標準ボタンを生成することが容易にできる。

この場合には、代表的なボタンマッチング法として広く実用に供されている動的計画法を利用した時間軸正規化マッチング法、いわゆるDPマッチング法(「日本音響学会誌、Vol. 27, No. 9, P. 483」)を適用することができる。

$K(m)$ が2以上の場合には、上記のようにして生成する単語標準ボタンは複数個となり単語マッチングの処理量が増加する。これに対し各音節毎に $K(m)$ 個の音節標準ボタンのうちから最も入力ボタンに類似していると思われるものをとり出してこれらを連結することにすれば、生成される単語標準ボタンは1つであり、処理量は大幅に減少できる。

このような方法としては、入力音声のうち、ある音節の存在する可能性のある区間を適当に切り出し、この区間で各音節標準ボタンとの類似度を求める方法いわゆる音節セグメンテーション(Segmentation)がある。

第2図は本発明における音節セグメンテーションの原理図である。第2図は、例として2音節長の単語「りく」の単語標準ボタンを生成する場合をとり、入力ボタンの区間 A と「り」の標準ボタンとの類似度および区間 B と「く」の標準ボタンとの類似度をそれぞれ求めればよいことを示している。このとき類似度として2つのボタンの包含関係を表わす値、例えば、特願昭60-153217に示される方法において用いられている値 Δ を用いることができる。

区間 A 、 B を求めるには、振巾ボタンの特徴を利用する方法や単純に入力区間を等分する方法など任意のものが可能である。第1図では、入力区間 $[S, E]$ を2つに等分する点 X を求め、適当な許容限度 ϵ を用いて、 $A = [S, X + \epsilon]$ $B = [X - \epsilon, E]$ として与えている。

(実施例)

次に図面を参照して本発明を詳細に説明する。

第1図は本発明の一実施例を示すブロック図である。第1図において、入力ボタンバッファ1の内容は単語辞書5の各単語とのボタンマッチングが行われる。音節セグメンテーション処理部2は、バッファ1の入力ボタンを単語辞書5に記憶されている単語を仮説して各音節に対応する区間を求める。音節標準ボタン選択部3では、処理部2で求められた各音節区間のボタンについて、標準ボタンバッファ6中の各音節標準ボタン(但し音節名が、仮説された単語内の対応する音節と一致するもの)との比較を行い最も類似している音節標準ボタンを連結標準ボタンバッファ7へ出力する。連結標準ボタンバッファ7には、選択部3で決定された音節標準ボタンを連

結したボタンが単語標準ボタンとして格納される。

単語マッチング部4は、ボタンバッファ1上の入力ボタンとボタンバッファ7上の単語標準ボタンとのマッチングを行い類似度を出力する。

以上の例において、セグメンテーション処理部2は、単純に音節数で等分したもの、単語に固有の情報（無声部を含むエネルギーの山がある等）を利用したもの等任意のものが使用可能である。標準ボタン選択部3も、入力ボタンの部分ボタンと標準ボタンの類似性を判定できるものであれば任意のものが可能であるが、効率の改善が本発明の利点であるので計算の容易なものである程よい。特願昭60-153217で述べられている方法はその一例である。単語マッチング部4も任意のものが可能であるが、DPマッチング法が認識精度の点で優れていることはよく知られており、この場合には最も適切である。

(発明の効果)

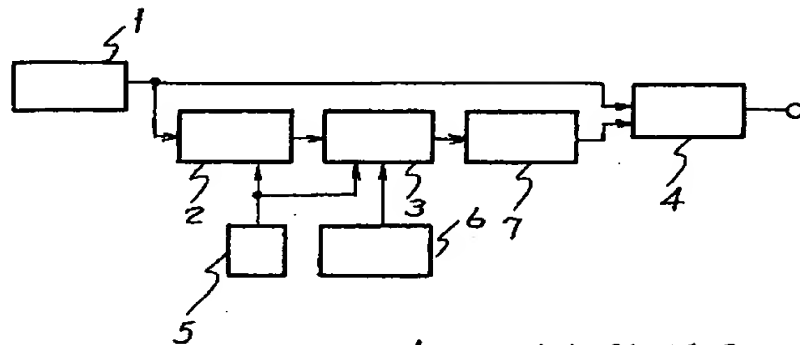
以上述べたように本発明によれば、複数のユニット標準ボタンを用いながら処理量の増加をもたらすことなくマッチングを行うことが可能となるので、認識精度と効率とを著しく改善したボタンマッチングを実現することができる音声認識におけるマッチング方法が実現できるという効果がある。

【図面の簡単な説明】

第1図は本発明の一実施例を示すブロック図、第2図は本発明における音節セグメンテーションの原理図である。

1……入力ボタンバッファ、2……セグメンテーション処理部、3……標準ボタン選択部、4……単語マッチング部、5……単語辞書、6……標準ボタンバッファ、7……連結標準ボタンバッファ。

【第1図】



- 1 ……入力ボタンバッファ
- 2 ……セグメンテーション処理部
- 3 ……標準ボタン選択部
- 4 ……単語マッチング部
- 5 ……単語辞書
- 6 ……標準ボタンバッファ
- 7 ……連結標準ボタンバッファ

【第2図】

